

Music Genres Reconsidered: Challenging Established Genres with a Data-driven Approach

Christian Hotz-Behofsits
WU Vienna
chotzbeh@wu.ac.at

Daniel Winkler
WU Vienna
dwinkler@wu.ac.at

Nils Wlömert
WU Vienna
nwloemert@wu.ac.at

Abstract

Consumers widely use music genres (e.g., pop, rock) for finding the right products. However, they are commonly arbitrary, not-standardized, disputed, and closely related genres often overlap. In this paper, we challenge established music genres (e.g., pop, rock) by comparing them to an entirely data-driven approach. To this end, we use a unique data set of revealed user preferences to carry out a context-based artist similarity. This measure is used in turn to find high-density artist clusters. The contribution of this paper is twofold. First, we investigate the differences between established music genres and data-driven clustering. Second, we provide implications for researchers and practitioners.

1. Introduction

Due to the hedonic nature of music, there is no objective measure for the fit between a given song and a given consumer. Every consumer has to experience a song to render their subjective judgment [1]. As an astronomical amount of new music is released each week, consumers need a way to filter out the content they will most likely not enjoy.

Nowadays, the majority of the revenue in the music industry is generated through streaming services such as Spotify or Apple Music [2]. This shift has given consumers access to millions of songs at a fixed monthly rate [3]. As part of the digital transformation, ever more artists can enter the market without relying on major labels for distribution [4]. However, most consumers will only enjoy listening to a (tiny) subset of available songs. Consequently, consumers rely on support to find the right products [5].

Researchers and practitioners particularly strive to reduce customers' search costs using machine learning (e.g., recommender systems), or bundling (e.g., playlists) [6, 7]. However, already a simple categorization of music based on its style (i.e., music genres) can alleviate the issue. While consumers are

typically unaware of the entire assortment, they are usually familiar with common music styles. Such categorizations are better known as *music genres* and include common labels such as *pop*, *rock*, or *punk*. These categories are important for evaluating the similarity between products [8]. In addition, while some (sub-)genres might be harder to distinguish, consumers can recognize the most popular genres within a fraction of a second [9].

The usage of such a simple categorization provides many benefits. First, as music genres have continually evolved with the industry [10], most people are already familiar with them. Second, genres are being used industry-wide across brick-and-mortar stores, download platforms, and streaming services. Notably, music genres have been an important concept used for production, distribution, and reception of music over the last decades [11]. Consequently, compared to recommender systems and playlists, music genres are ubiquitous and do not exhibit a vendor lock-in. Finally, music genres are not personalized. Therefore, music genres align with privacy and data protection laws and are not subject to biases commonly observed in machine learning (e.g., filter bubbles [12]).

Music genres allow consumers to easily find the music they are most likely to enjoy. However, music genres exhibit a significant shortcoming: Neither a ground truth nor a widely adopted standard exists [13, 14]. As a result, almost every player in the industry uses a different number of genres. Moreover, there is no clear opinion on whether each artist, album, or song exhibits a clear genre affiliation, how many genres, and which genres make sense. Even if different definitions often overlap considerably, it remains unclear to what extent they are appropriate.

In this paper, we challenge established music genres with an entirely data-driven clustering. To this end, we compare an artist clustering based on revealed customer preferences to established music genres. We provide guidance for which genres should be split up into sub-genres, removed, or combined.

2. Related Work

As music genres are strongly interwoven with the whole music industry, they have been used in different research fields. Besides sociology (e.g., [15]), they fuel marketing research (e.g., [16]), and music information retrieval (e.g., [17]).

In the context of information retrieval, the research community commonly treats music genres as a classification task [18]. Therefore, it is assumed that a ground truth exists [19, 20]. Genre-labeled examples are often provided in the course of contests such as *GTZAN* [17] and *ISMIR2004Genre* [21], or publications on audio similarity (i.e., *1517* [22]), and genre annotation (i.e., *tagtraum* [23]). Besides labels, the respective initiators also offer predictors (e.g., song features). Due to a lack of common standards, these tasks are confronted with three different challenges. First, different sets of genres are used. For example, *GTZAN* comprises ten (i.e., *blues*, *classical*, *country*, *disco*, *hip-hop*, *jazz*, *metal*, *pop*, *reggae*, and *rock*), *ISMIR2004Genre* six (i.e., *classical*, *electronic*, *jazz*, *metal.punk*, *rock.pop*, and *world*), and *1517* even 19 different genres. Secondly, the suitability of particular genres (e.g., *classic*) is questioned [15]. Finally, the provided labels exhibit quality issues (e.g., replicas, mislabeling, and distortions) [24]. Pálmason et al. [13] identified wrongly labeled examples in the ground truth using human raters (i.e., professional and semi-professional music protagonists) and updated them accordingly. However, these updates had only a minor impact on the classification error. Consequently, the authors conclude that a simple classification taxonomy is insufficient, and some content cannot be clearly labeled.

Moreover, the amount of categories is a double-edged sword. On the one hand, a higher number of genres increases the choosers' perceptions of assortment variety [25]. On the other hand, too much choice may backfire by increased cognitive effort [26].

However, there have been efforts to frame the problem as clustering instead of a classification task. Or at least to rethink music genre classification "from the ground up" [27, p.4]. In contrast to classification, Langenlois and Marques [28] propose an unsupervised clustering. In particular, their clustering is based on song features. Over time, similar systems have been introduced, and expanded that also analyze the content (e.g., [29, 30, 31]). Consequently, patterns emerge on the similarity of the content. However, it should be noted that content does not necessarily have to sound the same to be related [32]. For example, the anthems of football clubs can differ significantly in pitch, timbre,

or rhythm. To overcome this limitation, [33] propose to build the clustering on customer preferences instead of content. Customer preferences are taken from a survey. Specifically, they asked 1,523 people for three names of groups, singers, and artists they had consumed in the last month [34]. We extend this work by using observational data instead of the survey. This approach allows us to use significantly more observations and artist preferences per person to get a better picture of genres.

3. Analysis

In this section, we present a data-driven approach for categorizing music artists. The results are in turn compared with established music genres.

3.1. Data

Music genres Few artists can be categorized appropriately using a single main genre (e.g., *pop*). Consequently, music categorization can be very granular. It seems natural to describe a product using multiple genre names, adjectives and blend words in the interest of clarity. Examples include creations such as *blackened thrash*, *deep latin hip-hop*, or *chill synthwave*. In this vein, the music information services (e.g., last.fm, Chartmetric) distinguish between thousands of different "tags" (i.e., genre definitions). Besides quite specific labels (i.e., sub-genres), it also covers well-known main-genres such as *pop* and *rap*. We loaded the style tags (i.e., music genre) for 8,190 different artists via the Chartmetric API. As the aim is to analyze specific identifiers, we split combinations (e.g., *new chill synthwave*) into their tokens (e.g., *new*, *chill*, and *synthwave*). These parts are referred to as *genres*, *tags*, *tokens*, or *labels* in the following. Splitting combinations yields 1,511 different tokens.

Many of them are only used a few times. Consequently, we focus on the subset of the top 30. Since we want to have a global (i.e., country-independent) definition of genres, country-specific terms (e.g., *Swedish*, *German*, *Australian*, *UK*, and *Brazilian*) are removed. Figure 1 shows the frequency for the remaining tags. These resulting 25 tokens occur in more than 36% of all genre definitions.

Music preferences Over the last few years, we scraped 4,819,086 human-created collections of "loved songs" from a globally operating music streaming service. In this paper, we use only lists that contain songs by at least two different artists, such that they

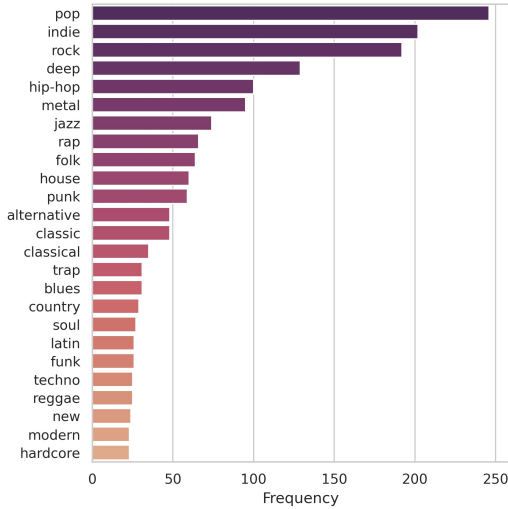


Figure 1. Frequency of top 25 tags.

can be used to learn context-based similarities. Enough training examples are required to train a machine learning model properly. Therefore, we only consider artists who occur at least 25 times in the training set. In total, this data set comprises 112,114 different co-liked artists.

3.2. Artist similarity

To create a similarity measure between artists, we use representational learning to map artists into a low dimensional (relative to the number of artists) space. In particular, we use the continuous bag-of-words (CBOW) neural network model proposed in [35]. The resulting vector representations are called embeddings and comprise semantic information (i.e., semantically similar artists are in each other's proximity). This unsupervised algorithm learns artist representations by predicting an artist based on the other artists liked by the same person. CBOW originates from text mining, where usually only the directly surrounding words are considered as predictive for the current word [36]. However, since we do not assume any influence of the sequential order of the likes in observed artist preferences, all other liked artists are used in our context.

The dimensionality of the vector space is a hyperparameter that is usually optimized. However, a ground truth of similarity between artists is not easily assessed and, if only feasible for popular ones [37]. Hence, we follow a rule of thumb. According to Google, the fourth root of the cardinality of the vocabulary is a good choice [38]. Based on our data set, this rule yields

a value of 19 (i.e., 18.3 rounded to the next integer).

3.3. Artist cluster

The challenges of music genres also pose a threat to music clustering. First, the lack of ground truth makes it an unsupervised clustering, which is not trivial to evaluate. Second, areas in the vector space with a high density of artists with low proximity should correspond to the same genre. Third, not every artist should necessarily fall into a specific genre [13]. We refer to the cases that cannot be easily clustered as noise. To this end, we use hierarchical density-based clustering [39]. It extends *Density-Based Spatial Clustering of Applications with Noise (DBSCAN)* such that it only requires one hyperparameter (i.e., *min_cluster_size*) [40]. Music genres are essential for the formation of music charts. Since most music charts feature at least the top 100 artists, we use this value as the minimum cluster size. HDBSCAN scales well to many observations, automatically determines a suitable number of clusters, and treats outliers (i.e., ambiguous artists) as noise. The algorithm thus accounts for the difficulties with music genres [13]. We use the artist representations learned by CBOW, a 19-dimensional real vector normalized to unit length, as input. Due to normalization, the Euclidean distance corresponds to the cosine similarity, which reduces computational complexity.

In HDBSCAN, the density around an artist is defined by how many others are in the immediate vicinity. This density is determined only for "core points", i.e., those located in higher density regions. An artist has to have a minimum number of other artists (*min_samples*) in its area to be considered as a core point. The higher the number of minimal samples, the fewer core points, fewer clusters, and more artists without cluster membership. By following best-practice, we set this hyperparameter equal to the minimum number of artists in a cluster (i.e., 100). The estimated model yields 14 different clusters.

3.4. Comparison

We assess similarity and association between genre i and cluster j based on two different measures: The *Szymkiewicz–Simpson coefficient (overlap)* and *normalized pointwise mutual information (NPMI)*.

Szymkiewicz–Simpson coefficient is a similarity measure that measures the overlap between two finite sets as follows [41]:

$$overlap(i, j) = \frac{c(i, j)}{\min(c(i), c(j))} \quad (1)$$



Figure 2. Data-driven artist clusters for a random subset of 2500 allocated artists (i.e., colors) and 500 noisy ones (i.e., grey).

Whereby, $c(i, j)$ is the number of artists labeled with genre i and are part of cluster j . Furthermore, we denote the number of artists labeled with genre i as $c(i)$ and the number of artists in cluster j as $c(j)$. Consequently, the overlap is a value in a range between 0 and 1.

Normalized pointwise mutual information is a measure of association that can be used for discrete random variables [42].

$$npmi(i, j) = \frac{\ln \frac{c(i, j)N}{c(i)c(j)}}{-\ln \frac{c(i, j)}{N}} \quad (2)$$

Whereby N denotes the total number of observations, $c(\cdot)$ occurrences along clusters or genres, and $c(\cdot, \cdot)$ common occurrences. The numerator comprises the pointwise mutual information, which we normalize by the self-information $-\ln \frac{c(i, j)}{N}$ to a range between -1 and 1.

While the chosen similarity measure (i.e., overlap) cannot take on negative values, npmi does.

4. Results

In the following, we will analyze the similarities (i.e., overlap) and associations (i.e., npmi) between existing genres and data-driven clusters. Specifically, we look at all 14 detected clusters and 24 genres. The results are presented in tabular form (see Table 1,

Table 2, and Table 3) and visualized as heat maps (see Figure 3, and Figure 4).

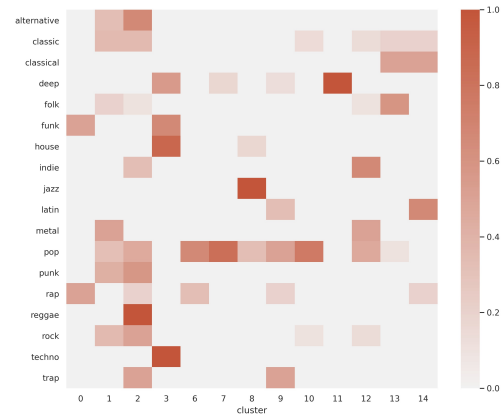


Figure 3. Overlap between data-driven clusters and established music genres.

Based on a comparison between the data-driven clustering and the human labels, we distinguish between three cases. First, we observe a clear overlap between a specific cluster and genre pair. In this case, the genre definition seems appropriate. Secondly, we may find clusters that do not exhibit any overlap with an existing genre. Consequently, the results indicate a concise cluster of artists which is not covered by genres. Finally, we expect some existing genres to be too generic, such that they are overlaps across multiple clusters. In such cases, we advise splitting the genre into sub-genres such

that there is a more apparent distinction.

First, we find no corresponding clusters for the tags *blues*, *country*, *hardcore*, *modern*, *new*, and *soul*. Thus, we do not consider them in visualizations. Moreover, we observe only weak cluster associations for the genres *punk*, *metal*, and the label *classic*. Both, the label *classic* and the term *classical* exhibit positive associations with cluster 13 and 14. However, these associations are more pronounced for *classical*. In addition, we also do not find any overlap at all for clusters 4 and 5. There seems to be no correspondence with established music genres here. This empirical evidence suggests to introduced additional tags to describe these styles of music.

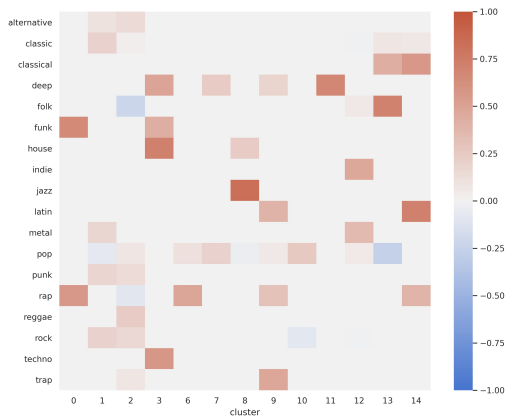


Figure 4. Association (i.e., npmi) between data-driven clusters and established music genres.

Second, we take a closer look at those instances which exhibit multiple associations across genres or clusters. In the case of genres, *pop* stands out because it overlaps with several clusters. However, this pattern is less pronounced in terms of associations, and some clusters are even negatively associated. Consequently, *pop* is not particularly useful in distinguishing music styles and should be split up into multiple sub-genres. In addition, the *rap* genre has a high association with multiple clusters and partly overlaps with *trap*, suggesting that some artists should be relabeled as the latter.

Third, it is also worth noting the term *deep*, which is mostly used as an additional specification of sub-genres. This label is the only one that shows substantial overlap with cluster 11. Moreover, its association is strongly positive. Therefore, *deep* seems to form its own style and could therefore be considered as a separate label. Such a spin-off could yield a more distinct grouping.

Finally, the genres *house* and *techno* are both strongly positively associated with cluster 3. Consequently, introducing a meta-genre for these

genre	cluster	overlap	npmi
trap	9	0.50	0.48
trap	2	0.50	0.08
rock	1	0.36	0.20
rock	2	0.50	0.16
rock	12	0.13	-0.01
rock	10	0.10	-0.08
rap	0	0.50	0.57
rap	6	0.33	0.49
rap	14	0.20	0.39
rap	9	0.20	0.31
rap	2	0.20	-0.09
punk	1	0.43	0.18
punk	2	0.57	0.14
pop	10	0.76	0.26
pop	7	0.83	0.20
pop	6	0.67	0.11
pop	2	0.45	0.07
pop	9	0.50	0.06
pop	12	0.47	0.05
pop	8	0.33	-0.04
pop	1	0.33	-0.07
pop	13	0.10	-0.26

Table 1. Overlap and npmi for trap, rock, punk, and pop

two should be considered. Given the overlap between genres, cluster 2 is prominent as it interferes with ten different tags simultaneously. However, this pattern is not strongly pronounced with regard to association, and it is even negative in two cases (i.e., *folk*, and *rap*).

Finally, we look at those genres for which we find clear correspondences between concrete clusters. This is especially the case for the genres *techno*, *indie*, and *jazz*. We observe a clear overlap with two clusters (i.e., 2 and 13) for *folk*, with cluster 2 having a negative association. In the case of *classical*, *funk*, *latin*, and *trap* there are overlaps with several clusters in each case, whereby one is always clearly more positively associated than the others.

5. Discussion and limitations

On the one hand, we do not find any corresponding cluster for the labels *blues*, *country*, *hardcore*, *modern*, *new*, and *soul*. Consequently, they could either be too ambiguous or specific. On the other hand, we find no genre matches for clusters 4 and 5. Consequently, one should avoid the problematic designations and find new ones for the clusters that cannot be assigned.

Both the terms *alternative* and *classic* show a weak association with the clusters found. For this reason, the

genre	cluster	overlap	npmi
techno	3	1.00	0.57
reggae	2	1.00	0.24
metal	12	0.50	0.36
metal	1	0.50	0.17
latin	14	0.67	0.71
latin	9	0.33	0.40
jazz	8	1.00	0.84
indie	12	0.67	0.48
indie	2	0.33	0.00
house	3	0.89	0.71
house	8	0.17	0.25

Table 2. Overlap and npmi for techno, reggae, metal, latin, jazz, indie, and house

genre	cluster	overlap	npmi
funk	0	0.50	0.66
funk	3	0.67	0.44
folk	13	0.60	0.71
folk	12	0.10	0.06
folk	1	0.20	0.00
folk	2	0.10	-0.22
deep	11	1.00	0.67
deep	3	0.56	0.49
deep	7	0.17	0.25
deep	9	0.12	0.19
classical	14	0.50	0.57
classical	13	0.50	0.44
classic	1	0.36	0.20
classic	13	0.20	0.08
classic	14	0.20	0.07
classic	2	0.36	0.03
classic	10	0.14	0.01
classic	12	0.13	-0.01
alternative	2	0.67	0.15
alternative	1	0.33	0.10

Table 3. Overlap and npmi for funk, folk, deep, classical, classic, and alternative

use of these labels should be reconsidered.

In contrast, the adjective *deep* does not only seem to function as an additive but is the only label that shows a strong positive association with cluster 11. For this reason, it should not only be used to define sub-genres but possibly be considered as a genre in its own right.

Interestingly, there is no clear cluster for *rock*. There are clear overlaps with clusters 1 and 2, but the corresponding associations are weakly positive. This could be due to the versatility of *rock*. For this reason, the genre could be too "broad" and not have any high-density regions.

Since pop is related to many clusters, providing additional labels for pop songs might improve the discoverability of these songs. In addition, there are both positive and negative associations to most clusters. However, these are all only weakly pronounced. Consequently, it is advisable to replace the label with more specific expressions.

Music genres are mainly used to categorize existing music. However, this classification also has at least an indirect (e.g., via charts) influence on the success of artists. So it is conceivable that this classification also influences the creation of music, customer preferences, and finally, consumption. Thus, an entirely data-driven approach would not be completely independent but would be influenced by existing definitions. In this paper, however, existing differences between data-driven approaches and existing categorizations are of interest. Thus, the paper aims not to create an entirely new categorization but to identify and show potential improvement in existing music genres.

In this work, we have chosen hyperparameters (i.e., *min_cluster_size*) such that the resulting clusters are large enough for music charting. One could also optimize this hyperparameter and perform a sensitivity analysis. However, this would go beyond the scope of this paper, so we leave it open for future research.

In addition, we did not offer designations for clusters without or with ambiguous genre agreement. On the one hand, this requires further investigation, and on the other hand, we lack the creativity to do so.

Furthermore, we have assumed that there is a genre definition at the artist level. Similar undertakings could be made at song or album level, and the results compared.

Moreover, the analysis at hand is based on user created "loved" collections of songs. Future research could use similar methodology based on other types of implicit feedback (i.e., consumption) or content of songs (e.g., lyrics, audio features), and investigate differences in the results. In addition, we assume that the clusters do not overlap. This assumption could be relaxed by

future research projects. In particular, applying fuzzy clustering may lead to clusters which represent the great variety of musical styles more accurately.

To extend the methodological approach, time-varying clusters could be used, allowing genres to change over time. Such an approach would account for music trends as observed in the industry [43]. Finally, this type of analysis can also be adopted for other products, such as movies and TV-shows, or books.

6. Summary

In this paper, we put established music genre definitions to the test. Music genres are a widely-used categorization of music that helps in the content search. Although they are used industry-wide, there is no unique definition. As a result, each provider uses its selection of genres, and the number also differs. To test established genre definitions (e.g., pop), we compare them with an entirely data-driven approach. Specifically, we create an artist similarity based on observed customer preferences. We use this metric in turn to obtain artist representations in a vector space. This representation allows us to detect areas with a high density of artists. These collections are subsequently referred to as clusters and can be seen as a data-driven alternative to music genres. For the genre definitions *jazz*, *techno*, and *reggae*, we find clear overlaps with data-driven clusters. The adjective *deep* seems to be suitable not only to indicate sub-genres (e.g., *deep house*) but also to exhibit a connection with a set of artists. The results also show the weaknesses of the *pop* genre. Although there are overlaps with clusters, they are ubiquitous and weak. We can also find overlaps with clusters in *rock* music, but these are only weakly associated. As a result, the genres of *pop* and *rock* should be reconsidered.

References

- [1] K. T. Lacher, "Hedonic consumption: Music as a product," *ACR North American Advances in Consumer Research*, vol. 16, pp. 367–373, 1989.
- [2] IFPI, "Global Music Report 2020," tech. rep., International Federation of the Phonographic Industry, 2020.
- [3] N. Wlömert and D. Papies, "On-demand streaming services and music industry revenues—Insights from Spotify's market entry," *International Journal of Research in Marketing*, vol. 33, no. 2, pp. 314–327, 2016.
- [4] L. Aguiar and J. Waldfogel, "Quality predictability and the welfare benefits from new products: Evidence from the digitization of recorded music," *Journal of Political Economy*, vol. 126, no. 2, pp. 492–524, 2018. Publisher: University of Chicago Press Chicago, IL.
- [5] A. Sela, J. Berger, and W. Liu, "Variety, vice, and virtue: How assortment size influences option choice," *Journal of Consumer Research*, vol. 35, no. 6, pp. 941–951, 2009. Publisher: The University of Chicago Press.
- [6] L. Aguiar and J. Waldfogel, "Platforms, promotion, and product discovery: Evidence from spotify playlists," tech. rep., National Bureau of Economic Research, 2018.
- [7] M. Schedl, P. Knees, B. McFee, D. Bogdanov, and M. Kaminskas, "Music recommender systems," in *Recommender systems handbook*, pp. 453–492, Springer, 2015.
- [8] A. Novello, M. F. McKinney, and A. Kohlrausch, "Perceptual evaluation of music similarity," in *ISMIR*, pp. 246–249, 2006.
- [9] R. O. Gjerdingen and D. Perrott, "Scanning the dial: The rapid recognition of music genres," *Journal of New Music Research*, vol. 37, no. 2, pp. 93–100, 2008. Publisher: Taylor & Francis.
- [10] M. Mauch, R. M. MacCallum, M. Levy, and A. M. Leroi, "The evolution of popular music: USA 1960–2010," *Royal Society open science*, vol. 2, no. 5, p. 150081, 2015. Publisher: The Royal Society Publishing.
- [11] E. Drott, "The end (s) of genre," *Journal of Music Theory*, vol. 57, no. 1, pp. 1–45, 2013. Publisher: Duke University Press.
- [12] A. Anderson, L. Maystre, I. Anderson, R. Mehrotra, and M. Lalmas, "Algorithmic effects on the diversity of consumption on spotify," in *Proceedings of The Web Conference 2020*, pp. 2155–2165, 2020.
- [13] H. Pálmason, B. P. Jónsson, M. Schedl, and P. Knees, "Music genre classification revisited: An in-depth examination guided by music experts," in *International Symposium on Computer Music Multidisciplinary Research*, pp. 49–62, Springer, 2017.
- [14] J.-J. Aucouturier and F. Pachet, "Representing musical genre: A state of the art," *Journal of new music research*, vol. 32, no. 1, pp. 83–93, 2003. Publisher: Taylor & Francis.
- [15] J. Vlegels and J. Lievens, "Music genres as historical artifacts: The case of classical music," *Connections*, vol. 35, no. 1, pp. 53–62, 2015.
- [16] J. Berger and G. Packard, "Are atypical things more popular?," *Psychological science*, vol. 29, no. 7, pp. 1178–1184, 2018.
- [17] G. Tzanetakis and P. Cook, "Musical genre classification of audio signals," *IEEE Transactions on speech and audio processing*, vol. 10, no. 5, pp. 293–302, 2002. Publisher: IEEE.
- [18] B. L. Sturm, "The state of the art ten years after a state of the art: Future research in music information retrieval," *Journal of New Music Research*, vol. 43, no. 2, pp. 147–172, 2014. Publisher: Taylor & Francis.
- [19] S. Lippens, J.-P. Martens, and T. De Mulder, "A comparison of human and automatic musical genre classification," in *2004 IEEE international conference on acoustics, speech, and signal processing*, vol. 4, pp. iv–iv, IEEE, 2004.
- [20] K. Seyerlehner, G. Widmer, and P. Knees, "A comparison of human, automatic and collaborative music genre classification and user centric evaluation of genre classification systems," in *International Workshop on Adaptive Multimedia Retrieval*, pp. 118–131, Springer, 2010.

- [21] P. Cano, E. Gómez, F. Gouyon, P. Herrera, M. Koppenberger, B. Ong, X. Serra, S. Streich, and N. Wack, "ISMIR 2004 audio description contest," *Music Technology Group of the Universitat Pompeu Fabra, Tech. Rep.*, 2006.
- [22] K. Seyerlehner, G. Widmer, and P. Knees, "Frame level audio similarity-a codebook approach," in *Proc. of the 11th Int. Conf. on Digital Audio Effects (DAFx-08)*, p. 31, 2008.
- [23] H. Schreiber, "Improving Genre Annotations for the Million Song Dataset.," in *ISMIR*, pp. 241–247, 2015.
- [24] B. L. Sturm, "An analysis of the GTZAN music genre dataset," in *Proceedings of the second international ACM workshop on Music information retrieval with user-centered and multimodal strategies*, pp. 7–12, 2012.
- [25] C. Mogilner, T. Rudnick, and S. S. Iyengar, "The mere categorization effect: How the presence of categories increases choosers' perceptions of assortment variety and outcome satisfaction," *Journal of consumer Research*, vol. 35, no. 2, pp. 202–215, 2008. Publisher: The University of Chicago Press.
- [26] J. T. Gourville and D. Soman, "Overchoice and assortment type: When and why variety backfires," *Marketing science*, vol. 24, no. 3, pp. 382–395, 2005. Publisher: INFORMS.
- [27] M. Lamont, "Looking back at Bourdieu," *Cultural Analysis and Bourdieu's legacy: settling accounts and developing alternatives*, pp. 128–41, 2010. Publisher: Routledge London.
- [28] T. Langlois and G. Marques, "Automatic Music Genre Classification Using a Hierarchical Clustering and a Language Model Approach," in *2009 First International Conference on Advances in Multimedia*, (Colmar, France), pp. 188–193, IEEE, July 2009.
- [29] W.-H. Tsai and D.-F. Bao, "Clustering music recordings based on genres," in *2010 International Conference on Information Science and Applications*, pp. 1–5, IEEE, 2010.
- [30] E. Tsunoo, G. Tzanetakis, N. Ono, and S. Sagayama, "Audio genre classification using percussive pattern clustering combined with timbral features," in *2009 IEEE International Conference on Multimedia and Expo*, pp. 382–385, IEEE, 2009.
- [31] S. S. Ghosal and I. Sarkar, "Novel Approach to Music Genre Classification using Clustering Augmented Learning Method (CALM).," in *AAAI Spring Symposium: Combining Machine Learning with Knowledge Engineering (I)*, 2020.
- [32] M. Schedl and P. Knees, "Context-based music similarity estimation," in *Welcome to the 3rd International Workshop on Learning Semantics of Audio Signals*, p. 59, Citeseer, 2009.
- [33] J. Vlegels and J. Lievens, "Music classification, genres, and taste patterns: A ground-up network analysis on the clustering of artist preferences," *Poetics*, vol. 60, pp. 76–89, 2017. Publisher: Elsevier.
- [34] J. Lievens, J. Siongers, and H. Waeye, *Participatie in Vlaanderen I: Basisgegevens van de Participatiesurvey*. Leuven: Acco, 2014.
- [35] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," in *Advances in neural information processing systems*, pp. 3111–3119, 2013.
- [36] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," *arXiv preprint arXiv:1301.3781*, 2013.
- [37] H. Ellis-Petersen, "Taylor Swift takes a stand over Spotify music royalties," *The Guardian*, Nov. 2014. Section: Music.
- [38] Google, "Introducing TensorFlow Feature Columns," Nov. 2017.
- [39] L. McInnes, J. Healy, and S. Astels, "hdbscan: Hierarchical density based clustering," *Journal of Open Source Software*, vol. 2, no. 11, p. 205, 2017.
- [40] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise.," in *Kdd*, vol. 96, pp. 226–231, 1996. Issue: 34.
- [41] M. K. Vijaymeena and K. Kavitha, "A survey on similarity measures in text mining," *Machine Learning and Applications: An International Journal*, vol. 3, no. 2, pp. 19–28, 2016.
- [42] G. Bouma, "Normalized (pointwise) mutual information in collocation extraction," *Proceedings of GSCL*, pp. 31–40, 2009.
- [43] K. Boughanmi, A. Ansari, and R. Kohli, "Dynamics of musical success: A Bayesian nonparametric approach," tech. rep., Technical report, Working paper, Columbia University, 2018.